

**Prediction of heart failure hospitalizations with wearable
activity monitors**

by

Ximin Li

A thesis submitted to The Johns Hopkins University in conformity with the
requirements for the degree of Master of Science.

Baltimore, Maryland

April, 2014

© Ximin Li 2014

All rights reserved

Abstract

Sixty one patients of the Advanced Cardiac Care Center of Columbia University Medical Center diagnosed with congestive heart failure (CHF) wore Actical, an accelerometer device that continuously recorded physical activity over a seven to nine months period. Over the course of the study, twenty two subjects were either hospitalized or had an emergency room visit. The goal of this thesis is to explore whether ambulatory monitoring of physical activity with accelerometers predicts clinically relevant adverse events in CHF patients. We introduce novel actigraphy summaries and identify prevalent patterns that explain roughly 60% of all CHF related episodes. These patterns can be detected as early as two months to three weeks prior to an episode.

Primary Reader: Vadim Zipunnikov, PhD

Secondary Reader: Jennifer Schrack, PhD

Acknowledgments

This thesis represents my study and work at Johns Hopkins Bloomberg School of Public Health, where I enjoyed wonderful time as a master student. I am grateful to my thesis advisor Dr. Vadim Zipunnikov, who introduced this interesting topic to me and was always being inspiring, patient and helpful along the way of my research. I would also like to thank Dr. Jennifer Schrack for being a member of my committee and reading my thesis. In addition, I would like to take this opportunity to thank all the professors and colleagues who taught me and helped me. Finally, I want to thank my parents who have always supported my study and life both physically and psychologically.

Contents

Abstract	ii
Acknowledgments	iii
List of Tables	vii
List of Figures	viii
1 Introduction	1
2 Heart Failure Study	4
2.1 Subjects	4
2.2 Actical	6
3 Time-Invariant Metrics	9
3.1 Log-transformation	9
3.2 Traditional Statistical Metrics	10
3.2.1 Mean	10

CONTENTS

3.2.2	Standard deviation	11
3.2.3	Skewness	11
3.2.4	Kurtosis	11
3.2.5	Mean over standard deviation ratio	12
3.3	Actigraphy Metrics	12
3.3.1	Active minutes and Active Mean	12
3.3.2	MaxH	13
3.3.3	MaxHMin	14
3.3.4	Total Variability	14
3.3.5	Fragmentation	15
3.3.6	Entropy	15
3.3.7	Example	16
4	Time of the day analysis	19
4.1	KS test	19
4.2	Day and Night	21
5	Fully Temporal Analysis	23
5.0.1	Principal Component Analysis	23
6	Results	25
7	Conclusion	29

CONTENTS

Vita	35
-------------	-----------

List of Tables

2.1	Demographic and Clinical Characteristics	7
2.2	Subjects that Changed Device	8

List of Figures

3.1	Subject 50. Fifteen activity features.	18
4.1	KS-test. P-value curves and the bootstrap intervals.	20
4.2	Subject 50: Day and Night.	21
5.1	Distribution of variance explained by the first ten PCs.	24
6.1	Subject 31. Pre-event dynamics of four summaries: Mean, Entropy, Frag(0), and Frag(500).	28
6.2	Subject 8. Dynamics of the eigenscore of PC2.	28

Chapter 1

Introduction

Heart failure (HF) is one of the leading chronic diseases in the elderly. There are approximately 5.1 million cases in United States, and the prevalence continues to rise [5]. The lifetime risk is 20% for those over age 40 in United States, and more than 650,000 new cases diagnosed annually [5]. Despite improved survival, the absolute mortality rate within five years of diagnosis approaches fifty percent[5]. The HF prevalence in Medicare-eligible population has increased from 90 to 121 per 1000 from 1994 to 2003, and this number is expected to rise to twenty percent in those older than 65 by 2050 [5]. Thus, the economic burden of HF is substantial and exceeds \$30 billion with more than half spent on hospitalization costs [5]. Therefore, identifying subjects with increased risk of hospitalization becomes important. Many risk models for HF hospitalizations use comorbidities such as atrial fibrillation, hypertension, diabetes mellitus, or other factors, such as physical fitness or level of income, as key

CHAPTER 1. INTRODUCTION

predictors. Decrease in physical function is one of the most common manifestations of HF that significantly influences quality of life. Therefore, a few standard measures of physical function such as the six minute walk or four hundred meter walk tests are usually included in risk models. However, emerging evidence indicates that the accuracy of risk models can be greatly improved by mobile technologies that allow real-time evaluation of a patient’s physical function.

Accelerometers, wearable sensors measuring acceleration, have been increasingly deployed in epidemiological studies to measure physical activity in the natural living environment. This mobile technology has quickly become a very popular option to explore complex associations between activity patterns and health outcomes [3, 1, 8, 12, 19, 13]. Compared to traditional activity assessment methods such as questionnaires and diaries [7, 14], accelerometers often provide an objective and more unbiased quantitative measurement of everyday physical activity. In addition, the use of accelerometers in studies has been shown to be a minimally invasive with no to little effect on subjects’ daily routines [10, 1, 18]. Thus, this technology provides an unprecedented opportunity to unveil previously hidden but pivotal physical activity signatures that manifest CHF.

To evaluate the clinical utility of continuous accelerometry in HF subjects, a prospective longitudinal cohort study was conducted by Columbia University Medical Center and GE Global Research. Subjects with clinical diagnosis of Heart Failure were enrolled in the study and followed for seven to nine months. Overall, sixty

CHAPTER 1. INTRODUCTION

one subjects were recruited to continuously wear an Actical device during the study period. The study coordinators visited subjects at their homes monthly and downloaded the data from the Actical as well as recorded any intercurrent events such as illness, hospitalization or emergency room visit. Actical can estimate daily energy expenditure using a validated algorithm. However, these estimates are often inaccurate and have a large variability in heterogeneous populations. Therefore, we will analyze physical activity using minute-by-minute raw data (activity counts) produced by the device.

In this thesis, we propose several novel summary statistics that extract information from raw data. Based on the summaries, we developed real-time physical activity profiles that track changes via within-day invariant characteristics of physical activity. We show that our summaries have a higher accuracy for predicting intercurrent events than traditional metrics. We also analyzed the data to determine the time of the day when the changes were the most prominent and incorporate this in our tracking models. Finally, we performed fully temporal analyses based on principal component.

The remainder of the thesis is organized as follows. Section 2 provides a short review of data collection and participant demographics. Section 3 describes the proposed summaries. Section 4 contains time-of-the-day adjusted analysis. Section 5 applies principal component analysis that takes into account within-day temporal dynamics. Section 6 concludes with a discussion of the results and future work.

Chapter 2

Heart Failure Study

2.1 Subjects

Columbia University Medical Center enrolled sixty one subjects with a clinical diagnosis of Heart Failure. The subjects agreed to continuously wear an Actical physical activity monitoring device during the study period, which varied from subject to subject and ranged from seven to nine months. The study coordinators visited subjects at their homes monthly and uploaded the data from the devices as well as recorded any intercurrent events such as illness, hospitalization or emergency room visit. Actical devices were set to record activity counts for every minute of each day, resulting in 1440 measurements per day.

To ensure enough data for our analysis, we excluded subjects who had less than sixty days of data overall as well as subjects who had an intercurrent event within first

CHAPTER 2. HEART FAILURE STUDY

thirty days of the study. As a result, forty-eight subjects were studied. We have split all subjects into two groups: those with no registered events (23 subjects) and those with registered events (25 subjects). In the event group, one subject died, seven visited the emergency room, nine were hospitalized, five experienced intercurrent illnesses, and three underwent outpatient procedures.

Table 2.1 compares the two groups at baseline across different demographic and clinical characteristics. The average age for the event group is older than that of the non-event group, but the difference is not significant. The two groups are neither different in gender or race. The non-event group has a greater proportion of subjects maintaining regular exercise. The event group is not significantly lower in functional capacity at baseline: the 6 minute walk test distance is non-significantly shorter for the event group. The average body mass index and the ejection fraction are about the same in both groups. In addition to demographics, daily summary statistics for the first week of the study were calculated for each subject and the medians were compared. None of the summary statistics were statistically different. More details on summaries and their definitions are provided in Section 3. We also split the day into two non-overlapping twelve-hour intervals which we refer to as Day (6am-6pm) and Night (6pm-6am). There were no statistically different summaries between Day and Night. Therefore, we concluded that at the baseline two groups seem to be similar across the wide range of demographic and clinical characteristics.

2.2 Actical

The study used an Actical activity monitor, a small wristwatch-like device that can provide metabolic equivalents (METs)-a measure of oxygen consumption. However, we analyzed the raw data (activity counts) recorded by the omnidirectional accelerometer in the device at minute-by-minute epochs. Physical activity is a strong predictor of health [20, 21]. Previous uses of actigraphy for health monitoring has centered on sleep disorders [2, 17] and population level differences in physical activity across large epidemiological studies [19, 16].

There were seven subjects that have changed the device during the study. Table 2.2 reports the subject IDs and the week when subjects changed the device. Among the seven subjects, participants 4, 6, and 19 showed a significant change in the activity levels before and after the device change. In order to maintain consistency in the data, we excluded these subjects when we performed comparisons in Section 4 using Kolmogorov-Smirnov test.

There were six subjects with overlapping periods adjacent to the data upload visit. For cases like these, we resolve the discrepancy by using the activity data from the period prior to the upload.

CHAPTER 2. HEART FAILURE STUDY

Table 2.1: Demographic and Clinical Characteristics

Table 1. Demographic and Clinical Characteristics				
	Total (N=48)	Event Group (N=25)	None Event Group (N=23)	
Characteristics	Mean (SD)	Mean (SD)	Mean (SD)	P Value
Age	61.46 (12.78)	63.64 (12.68)	59.09 (12.73)	0.2037
Male sex,no	23	13	10	
Body mass index(kg/m^2)	28.79 (6.41)	29.36 (7.18)	28.17 (5.56)	0.7799
6 minute walk test	391.54 (113.44)	378.6 (120.93)	405.61 (105.56)	0.4764
Race,no.				
White	31	17	14	
Black	10	5	5	
Other	7	3	4	
Regular exercise	15	5	10	
Ejection fraction, %	45.35 (14.76)	45.8 (15.86)	44.87 (13.80)	1
BNP1, %	252.78 (388.79)	212.14 (189.87)	297.28 (530.42)	0.6809
BNP2, %	227.78 (187.70)	268.1 (368.12)	207.63 (109.20)	1
Diabetes, no.	12	5	7	
Hypertension, no.	27	13	14	
Ischemic,no	8	5	3	
Anemia,no.	11	7	4	
1st week daily summary median of activity count				
Mean	3.10 (0.53)	3.03 (0.55)	3.17 (0.51)	0.6675
actmin	828.35 (5.11)	799.12 (110.16)	860.13 (104.75)	0.1373
active mean	5.11 (0.38)	5.08 (0.38)	5.15 (0.37)	0.7435
M6M	7.13 (0.46)	7.06 (0.54)	7.19 (0.34)	0.325
2M6M	7.07 (0.46)	7.01 (0.55)	7.14 (0.34)	0.461
Fragmentation-thredhold 0	232.67 (47.54)	218 (58.38)	235.5 (40.83)	0.8365
Fragmentation-thredhold 100	262.20 (45.98)	251.2 (42.52)	264.30 (54.17)	0.7647
Fragmentation-thredhold 500	157.49 (59.34)	147.24 (57.33)	161.37 (60.03)	0.8124
1st week daily summary median of activity count-Day				
Mean	3.96 (0.75)	3.85 (0.82)	4.08 (0.66)	0.461
active mean	5.24 (0.42)	5.20 (0.41)	5.30 (0.43)	0.5806
M6M	7.07 (0.47)	7.00 (0.55)	7.14 (0.36)	0.3563
Fragmentation-threshold 0	103.31 (30.83)	103.9 (31.23)	102.67 (31.08)	0.926
Fragmentation-threshold 100	147.40 (31.90)	142.2 (25.91)	147.5 (39.95)	0.5289
Fragmentation-threshold 500	105.26 (42.18)	96.48 (40.81)	109.54 (43.76)	0.8526
1st week daily summary median of activity count-Night				
Mean	2.23 (0.56)	2.19 (0.59)	2.27 (0.53)	0.461
active mean	5.24 (0.42)	5.19 (0.41)	5.30 (0.43)	0.5806
M6M	6.70 (0.44)	6.65 (0.50)	6.76 (0.37)	0.7281
Fragmentation-threshold 0	128.75 (27)	127.72 (31.97)	129.87 (20.97)	0.8607
Fragmentation-threshold 100	117.25 (25.44)	114.8 (25.03)	115.74 (26.44)	0.6349
Fragmentation-threshold 500	50.58 (23.96)	49.48 (24.94)	51.78 (23.36)	0.7333

CHAPTER 2. HEART FAILURE STUDY

Table 2.2: Subjects that Changed Device

Subject ID	Week
4	24
6	34
19	10
45	36
46	7
54	12
61	9

Chapter 3

Time-Invariant Metrics

In this section, we put forward novel actigraphy metrics that can simultaneously characterize multiple features of physical activity trajectories over the course of a day. We will demonstrate that it will greatly help with modelling longitudinal trajectories for HF events.

3.1 Log-transformation

Depending on physical activity levels, minute-by-minute activity counts can take any value between zero and six thousand. The distribution of activity counts were highly skewed to the right. To make the distribution more symmetric, we took the natural logarithm of each minute-by-minute activity count. Let $y_{ij}(t)$ denote the minute activity count for subject i at day j at minute t , then the log-transformed

CHAPTER 3. TIME-INVARIANT METRICS

count is $y'_{ij}(t) = \log(y_{ij}(t) + 1)$. After the transformation, zero activity count is transformed to zero with non-zero activity counts is rescaled to a much smaller range with a symmetric distribution.

3.2 Traditional Statistical Metrics

We start with reviewing traditional moment based statistical summaries. To study the pattern of subject activity over time, we rearrange y'_{ij} into a $1440 \times N_i$ matrix, where N_i denotes the total number of days for subject i .

3.2.1 Mean

Mean is the most commonly used summary statistics that measures the central value of a probability distribution or a data set. After the log transformation, the distribution of activity count is approximately normal, therefore, the mean should approximate well the center of the distribution. To evaluate the central tendency, the daily mean activity count is calculated as

$$M_{ij} = \frac{\sum_{t=1}^{1440} y'_{ij}}{1440}. \quad (3.1)$$

3.2.2 Standard deviation

Mean and variance is widely used to describe any distribution. In order to be consistent with the units of the mean, we calculate the square root of the variance, standard deviation, for subject i at day j as:

$$SD_{ij} = \sqrt{\frac{\sum_{t=1}^{1440} (y'_{ij}(t) - M_{ij})^2}{1399}}. \quad (3.2)$$

3.2.3 Skewness

Skewness is used to measure the deviation from the symmetry. We calculate it as

$$Skewness_{ij} = \frac{1}{1440} \sum_{t=1}^{1440} \left(\frac{y'_{ij}(t) - M_{ij}}{SD_{ij}} \right)^3 \quad (3.3)$$

3.2.4 Kurtosis

Kurtosis is used to measure the shape of the distribution and concentration around the mean:

$$Kurtosis_{ij} = \frac{1}{1440} \sum_{t=1}^{1440} \left(\frac{y'_{ij}(t) - M_{ij}}{SD_{ij}} \right)^4 \quad (3.4)$$

3.2.5 Mean over standard deviation ratio

Finally, the mean over the standard deviation ratio is used as a normalized measure that takes into account variability around the mean and is defined as

$$MoSD_{ij} = \frac{M_{ij}}{SD_{ij}}. \quad (3.5)$$

3.3 Actigraphy Metrics

When we collapse ultra dense actigraphy signal in just a few moment-based summaries we unavoidably suffer from enormous loss of information. However, this can be partially compensated through the introduction of interpretable actigraphy metrics that can simultaneously characterize multiple features of physical activity trajectories over the course of a day. Although, these metrics can be calculated, say, within an hour or a bout of activity, we will explore them at daily level.

3.3.1 Active minutes and Active Mean

The first metric is the total number of active minutes. This metric will differentiate between situations when a subject has a high total daily mean accumulated over relatively long vs short periods of activity. Let $I_{ij}(t, h)$ denote whether activity count for subject i at day j at minute t is larger than a threshold h , that is $I_{ij}(t, h) = 1$ if $y_{ij}(t) > h$ and $I_{ij}(t, 0) = 0$ if $y_{ij}(t, h) < h$. Then the total number of active minutes

CHAPTER 3. TIME-INVARIANT METRICS

can be written as

$$ActMin_{ij}(h) = \sum_{t=1}^{1440} I_{ij}(t, h). \quad (3.6)$$

Active mean is the total daily activity count divided by total number of active minutes. It could be viewed as an average level of activity given that a subject is active. Please note that while means take into account nonactive minutes, $ActM_{ij}(h)$ summarizes the activity counts only over active minutes:

$$ActM_{ij}(h) = \frac{1440M_{ij}}{ActMin_{ij}(h)}. \quad (3.7)$$

3.3.2 MaxH

$MaxH$ is the mean activity count of the H minutes with highest activity. In other words, if we order all activity counts as $\tilde{y}'_{ij}(1) \geq \tilde{y}'_{ij}(2) \geq \dots \geq \tilde{y}'_{ij}(1440)$, then $MaxH$ is defined as

$$MaxH_{ij} = \frac{\sum_{k=1}^H \tilde{y}'_{ij}(k)}{H}. \quad (3.8)$$

In the analysis below we use $Max10$, however, H could be adapted to the subject using a pre-defined objective criteria.

3.3.3 MaxHMin

MaxHMin count is the mean activity count in the H minute time interval with the highest cumulative activity count. [11] showed that *Max6Min* is highly correlated with 6 minute walk test (6MWT) which is widely used clinically to evaluate the functional capacity. *MaxHMin* can be defined as

$$MaxHMin_{ij} = \max_{t_0} \frac{\sum_{t=t_0+1}^{t_0+H} y'_{ij}}{H}. \quad (3.9)$$

In our analysis, we will [11] and use *Max6Min*. However, we will also explore how sensitive the results are to the choice of the time window, H .

3.3.4 Total Variability

The total variability is the cumulative absolute difference in activity counts between two adjacent minutes. It can be calculated as

$$TotalVar_{ij} = \sum_{t=2}^{1440} |y'_{ij}(t) - y'_{ij}(t-1)|. \quad (3.10)$$

It can be thought of as a numerical version of the functional total variability. A more general version is

$$TotalVar_{ij}(p) = \sum_{t=1+p}^{1440} |y'_{ij}(t) - y'_{ij}(t-p)|. \quad (3.11)$$

3.3.5 Fragmentation

In the free-living environment each activity has its own pattern. Some activities are moderate and do not last long, such as opening a door; some are short but intense, such as short distance running; some are moderate but last a while, such as cooking; and some are intense and last long such as exercising. A fragmentation index below keeps track of the number of times when activity status changed from inactive to active or vice versa from active to inactive. Fragmentation index can be defined as

$$Frag_{ij}(h) = \sum_{t=2}^{1440} |I_{ij}(t, h) - I_{ij}(t-1, h)|. \quad (3.12)$$

Similarly to the total variability, a more general version of fragmentation index can be defined as

$$Frag_{ij}(p, h) = \sum_{t=1+p}^{1440} |I_{ij}(t, h) - I_{ij}(t-p, h)|. \quad (3.13)$$

We will explore, fragmentation with h equal to 0, 100, 500, and 1000.

3.3.6 Entropy

Our last metrics is entropy. Entropy is a popular measure that quantifies the uncertainty of a random variable [9]. For actigraphy data, we can assume a finite discrete probability distribution p_k such that $\sum_{k=0}^K p_k = 1$. Then, entropy can be

CHAPTER 3. TIME-INVARIANT METRICS

calculated as

$$Entropy_{ij} = - \sum_{k=0}^K p_k \log(p_k), \quad (3.14)$$

where p_k is estimated using maximum likelihood approach applied to data representing day j in subject i .

3.3.7 Example

To further illustrate usefulness of derived features, we show trends in and interactions between fifteen different derived activity features plotted in Figure 3.1 for one of the CHF patients with a registered hospitalization. The subject has information for 39 weeks, and each weekly boxplot represents the distribution of seven daily summaries within a certain week. The blue vertical line shows the subject underwent a hospitalization at week 27. While the daily summary statistics go up and down over time, we see a sudden change at the week that event happened, especially for the mean, standard deviation, Max6Min, Frag, and TotalVar. The average daily mean activity count reached a high level the week before the event, then suddenly dropped down to a low level during the event week. Starting from week 28, we can track the course of the recovery that could be quite informative as well. One interesting phenomenon is that the Fragmentation and Total Variability show an opposite change to mean and standard deviation: while the mean and standard deviation suddenly dropped at the event week, the Fragmentation and Total Variance suddenly increased.

CHAPTER 3. TIME-INVARIANT METRICS

Thus, although the total activity level decreased, the subject days became much more fragmented.

As we discuss in Section 6, we have identified the most typical patterns and these patterns have explained roughly 60% of all hospitalization episodes and can be detected as early as two months to three week prior to an episode.

CHAPTER 3. TIME-INVARIANT METRICS

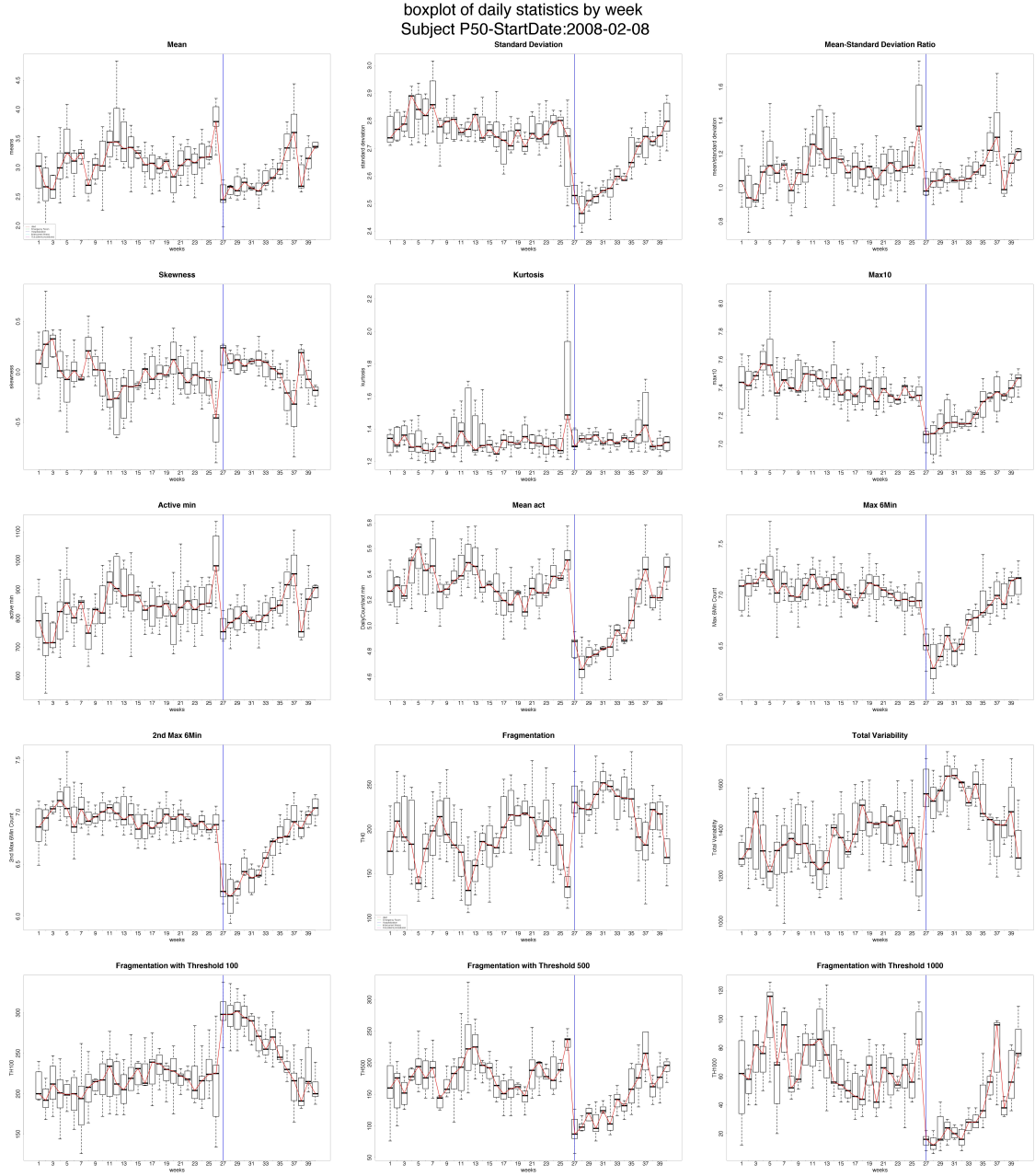


Figure 3.1: Subject 50. Fifteen activity features.

Chapter 4

Time of the day analysis

In the previous section, all metrics were time-invariant. Therefore, there is a loss of information about the temporal distribution of activity over the course of a day. In this section, we explore whether the activity distribution is different at different times of a day between the two groups. We will use the Kolmogorov-Smirnov (KS) test, a nonparametric test requiring minimal assumptions.

4.1 KS test

We start with comparing the distributions of the summary statistics hour by hour between the event group and non-event group. Particularly, we analyze ActMin, ActM, Max6Min, Frag0, Frag100, and Frag500 which have been calculated for each hour. To estimate uncertainty related to relatively small sample sizes, we sub-sampled

CHAPTER 4. TIME OF THE DAY ANALYSIS

fifteen subjects from each group, and hour-by-hour distributions of the metrics have been compared by KS-test. The bootstrap procedure has been repeated 100 times. Figure 4.1 reports the smoothed medians as well as 80% and 90% bootstrapped confidence intervals. Please note that we have applied the Bonferroni correction for the total number of tests report both unadjusted, 0.05, and adjusted, $\frac{0.05}{24}$, significance levels.

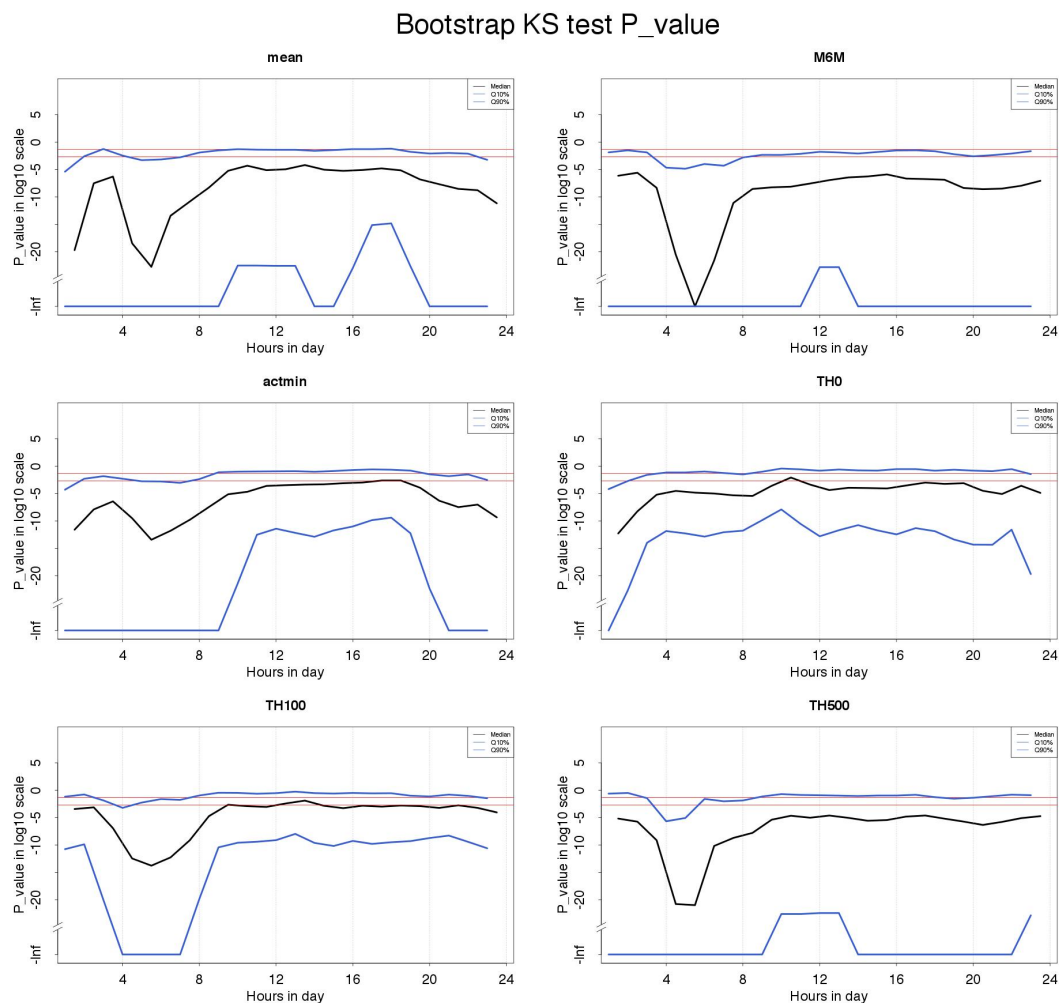


Figure 4.1: KS-test. P-value curves and the bootstrap intervals.

CHAPTER 4. TIME OF THE DAY ANALYSIS

From the graph, we see that the groups are quite different. All medians fell well below the adjusted significance level. We can also see that the difference for the event and non-event groups is the most significant for periods 10pm-2am and 3am-7am.

4.2 Day and Night

As shown above, the event and non-event group are the most different in evening and night hours. Therefore, we separate the day into two time intervals: Day, which we define as 6am-6pm and Night which we define as 6pm-6am. Please note that these intervals use information from two adjacent days.

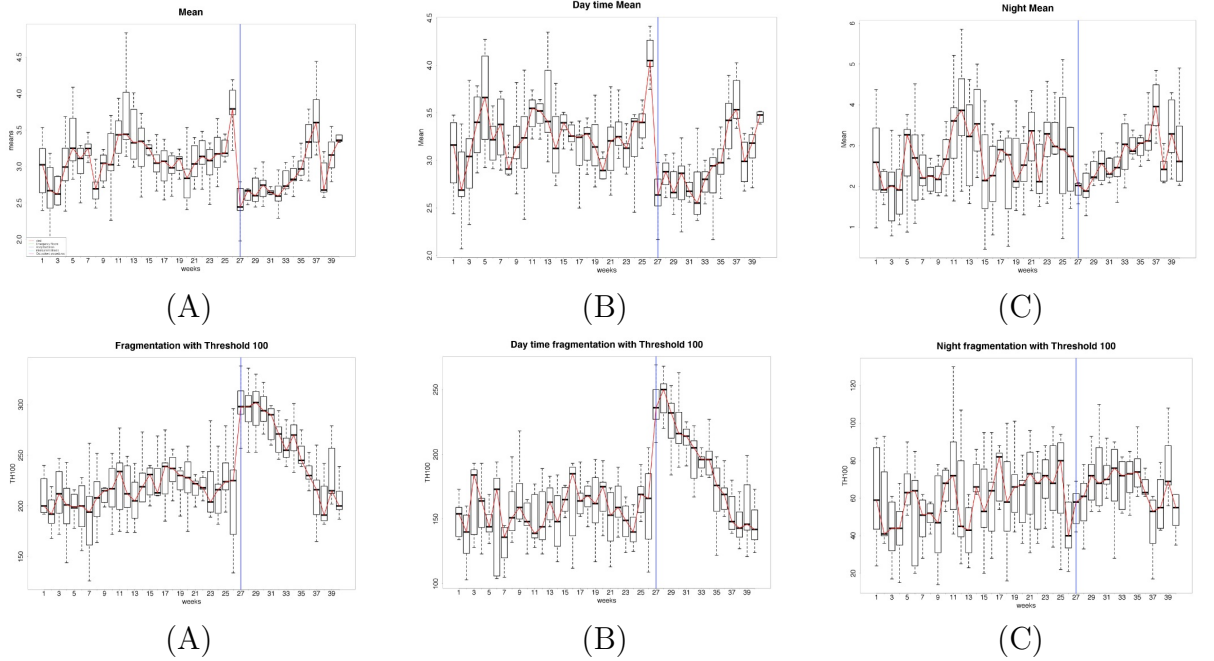


Figure 4.2: Subject 50: Day and Night.

In Figure 4.2, both Day and Night has similar trends as the whole day summaries.

CHAPTER 4. TIME OF THE DAY ANALYSIS

However, Night exhibits a less pronounced pattern.

Chapter 5

Fully Temporal Analysis

The goal of this section is to adapt principal component analysis (PCA) to accommodate information about the temporal distribution of activity over the course of a day.

5.0.1 Principal Component Analysis

Let us denote $Y_{ij} = [y_{ij}(1), y_{ij}(2), \dots, y_{ij}(1440)] \in R^{1440 \times 1}$, and $\mathbf{Y}_i = [Y_{i1}, Y_{i2}, \dots, Y_{iN_i}] \in R^{1440 \times N_i}$, where N_i is the number of days subject i in the study.

Principal components analysis (PCA) is a very popular dimension reduction technique. At the first step, PCA de-means vectors \mathbf{Y}_{ij} 's representing daily activity profile by subtracting the sample average, $\boldsymbol{\mu} = \sum_{i,j} \mathbf{Y}_{ij} / n$, where $n = \sum_{i=1}^J N_i$ and J is the sample size. We will use the same notation, \mathbf{Y}_{ij} , for the demeaned vectors. Second step of PCA calculates sample covariance matrix $\hat{\mathbf{K}}$ as $\hat{\mathbf{K}} = \frac{1}{n} \sum_{i,j} \mathbf{Y}_{ij} \mathbf{Y}_{ij}^T$. Note that

CHAPTER 5. FULLY TEMPORAL ANALYSIS

if the singular value decomposition of the demeaned data matrix $\mathbf{Y} = \mathbf{V}\mathbf{\Sigma}\mathbf{U}'$, then the covariance operator $\hat{\mathbf{K}}$ can be decomposed as $(1/n)\mathbf{V}\mathbf{\Sigma}^2\mathbf{V}'$, where orthonormal columns of the $p \times r$ matrix \mathbf{V} are eigenvectors and non-negative diagonal elements $\sigma_1^2/n \geq \sigma_2^2/n \geq \dots \geq \sigma_r^2/n > 0$ are eigenvalues. Third step of PCA calculates the n -dimensional vectors of eigenscores are calculated as $\boldsymbol{\xi}_{ij} = \mathbf{V}'\mathbf{Y}_{ij} = (\mathbf{\Sigma}\mathbf{U}')_{ij}$. In practice, only a few, say, K leading eigenvectors are used to parsimoniously model original images as $\mathbf{Y}_i = \boldsymbol{\mu} + \sum_{k=1}^K \mathbf{v}_k \xi_{ik}$. The number of eigenvectors, K , is typically chosen to make the proportion of variance explained, $p(k) = (\sigma_1^2 + \dots + \sigma_k^2)/(\sigma_1^2 + \dots + \sigma_r^2)$, large enough. Formal rank-selection approaches are based either on linear mixed effects models as in [4] or Bayesian models as in [6] and [15]. In our application, we choose 10 leading principal components. The proportion of variance explained, $p(10)$, varied between 0.55 and 0.9 with mean 0.77 and standard deviation 0.6. Figure 5.1 shows the distribution of $p(10)$ across all subjects.

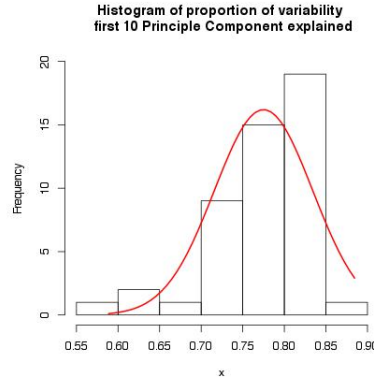


Figure 5.1: Distribution of variance explained by the first ten PCs.

Chapter 6

Results

Using metrics proposed in Section 3, we have developed comprehensive real-time physical activity profiles that allow to track and explore longitudinal changes during the course of the study. Appendix contains the profiles for all subjects. For completeness we created profiles for all available subjects. Each subject represented by a seven-page profile. To visually explore dynamics of change over 8 to 10 months of data, our approach is to summarize each day with the traditional and actigraphy metrics. Each of the metrics represents different aspect of the physical activity pattern observed in this day. Observing simultaneous dynamics of all the metrics prior to the event, we enormously reduce dimension and greatly increase chances of discovering informative patterns that emerging before the event. Each week is represented as a boxplot that serves as a visual smoother providing information both about the weekly mean of a summary and variability around the mean. First two pages of a

CHAPTER 6. RESULTS

profile report M_{ij} , $ActM_{ij}$, $Max6Min_{ij}$, $Frag_{ij}(0)$, $Frag_{ij}(100)$, and $Frag_{ij}(500)$, for the entire day (first column) as well as for Day (6am-6pm, middle column) and Night (6pm-6am, right column). Next page reports SD_{ij} , $MoSD_{ij}$, $Skewness_{ij}$, $Kurtosis_{ij}$, $Max10_{ij}$, $ActMin_{ij}$, second $Max6Min_{ij}$, $TotalVar_{ij}$, $Frag_{ij}(1000)$, and $Entropy_{ij}$. Page four of a profile reports $MaxHMin$ with window size H ranging from one to ten minutes of uninterrupted activity. The purpose of this analysis is to determine an optimal within-day window size that can be subject-specific. Pages five and six report the results of the principal component analysis with page five showing longitudinal dynamics of the first ten principal components and page six showing the components themselves. Finally, page seven shows a heatmap of activity counts with each row representing two adjacent days. The week of the event is marked with a vertical line where the color corresponds to the type of the event: died(red), emergency room (green), hospitalization (blue), intercurrent illness (cyan), outpatient procedures (pink).

Although, there is no single metric that consistently revealed a pre-event trend, we identified several metrics that are the most informative in capturing symptomatic pre-event changes. $ActM$ (Activity Minutes) captured patterns for subjects 10, 17, 39, 47, 49, 50, and 59. Night M (Mean) captured patterns for subject 10, 17, 39, 47, 49, 50, and 60. Night $Frag(0)$ captured patterns in subjects 9, 31, 39, 42, 49, 50, and 9. $Frag(500)$ captured patterns in subjects 9, 16, 24, 25, 31, 42, and 50. Day $ActM$ (Active mean) captured patterns in subjects 24, 25, 31, 42, 47, and 50.

CHAPTER 6. RESULTS

Frag(0) captured patterns in subjects 9, 31, 39, 49, 50, and 59. *Frag(100)* captured patterns in subjects 8, 9, 31, 37, 42, and 50. *Frag(100)* captured patterns in subjects 8, 9, 31, 37, 42, and 50. *Entropy* captured patterns in subjects 9, 15, 17, 31, 47, and 50.

Figure 6.1 illustrates how we can track dynamics of the proposed summaries. It reports pre-event period for subject 31 who visited emergency room during week 22. Four two plot panels show four summaries: Mean, Entropy, *Frag(0)* and *Frag(500)*. We use first two weeks as a benchmark. The weeks after the first two are formally compared with the benchmark with KS, Mann-Whitney-Wilcox, and Siegel-Tukey test. Bottom plot of each of the four panels at Figure 6.1 reports p-value curves for the tests. Both pre-event weeks 20 and 21 are significantly different from the benchmark with the downward trend during the event week.

Fully-temporal analysis using principal components revealed patterns in some of the subjects where our time invariant metrics have not shown any significant trends. For example, in Subject 8, the eigenscore corresponding to the second principal component peaks exactly at the event-week with emergency room visit (see Figure 6.2). This finding is consistent with the pattern observed on Day M (Mean) plot.

Overall, our approach revealed patterns in 17 out of 25 subjects in the event group. These patterns include a sudden change in the average level of a particular summary or a consistently increasing or decreasing trend in weeks preceding to the event. The observed patterns emerge as early as two months to three weeks before the event.

CHAPTER 6. RESULTS

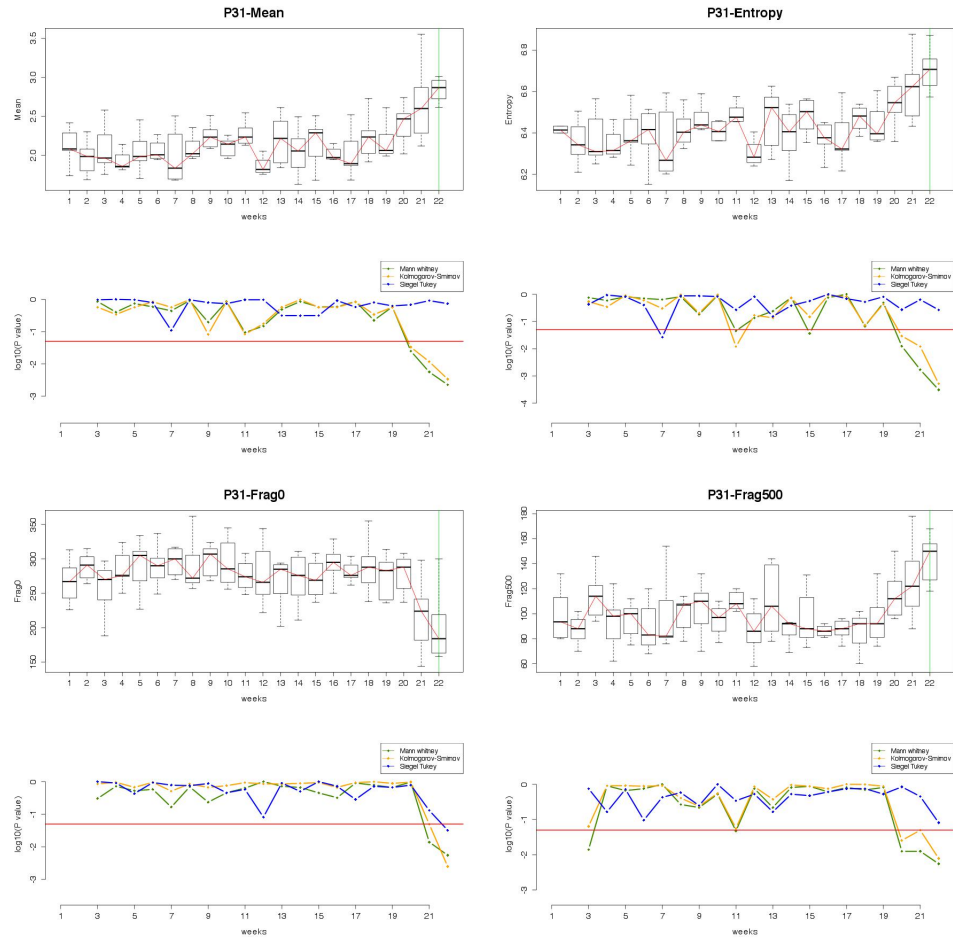


Figure 6.1: Subject 31. Pre-event dynamics of four summaries: Mean, Entropy, Frag(0), and Frag(500).

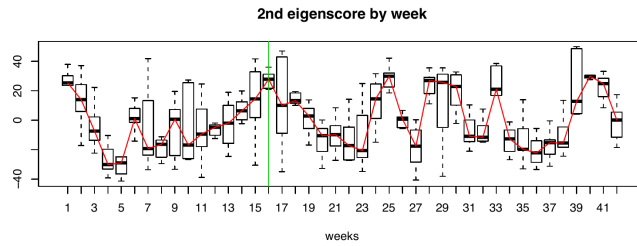


Figure 6.2: Subject 8. Dynamics of the eigenscore of PC2.

Chapter 7

Conclusion

Technological advances have made many wearable devices to monitor physical activity and physiology available for use in large epidemiological studies. These devices produce enormous volumes of continuously monitored actigraphy data which require vast amounts of database storage, cost millions of dollars, and require years of follow up. However, the scale of the data and the heterogenous nature of physical behavior across days and populations of subjects make statistical analysis and inference about cause and effect extremely challenging.

In this thesis, we proposed several novel actigraphy metrics that extract information from raw data. Based on these within-day invariant characteristics of physical activity, we developed real-time physical activity profiles that allow to track longitudinal changes. We showed that our summaries are highly informative for identifying changes in physical activity preceding intercurrent events such as hospitalizations and

CHAPTER 7. CONCLUSION

emergency room visits. We also analyzed the data using the fully temporal analysis based on principal components which provided important additional information and revealed patterns that have not been detected by time-invariant metrics. Although, this work demonstrates a considerable potential of using actigraphy for ambulatory monitoring of CHF patients, building formal risk prediction model is remained to be done.

Our metrics can be seamlessly incorporated into analytical frameworks that specifically target continuously monitored physical activity. For future research, we see a few possible extensions. Using our metrics, researchers can i) characterize intra- and inter-day connections among a range of physical activity features and explore associations with health outcomes, ii) analyse how principal activity patterns interact with each other, and iii) develop models that can both capture general interactions between activity features and make hypothesis-driven predictions for future actigraphy studies. These methods will make it possible to relate complex actigraphy data to health responses of interest for statistically rigorous outcomes-based research and clinical prediction, and create a reproducible data analytical pipeline for discovering, investigating and quantifying heterogeneous activity phenotypes in a robust, scalable and generalizable fashion.

Bibliography

- [1] A. A. Atienza and A. C. King. Comparing self-reported versus objectively measured physical activity behavior: A preliminary investigation of older filipino american women. *Research quarterly for exercise and sport*, 76(3):358–362, 2005.
- [2] Colleen E. Carney, Laura E. Lajos, and William F. Waters. Wrist actigraph versus self-report in normal sleepers: Sleep schedule adherence and self-report validity. *Behavioral Sleep Medicine*, 2(3):134–143, 2004.
- [3] K. M. Culhane, M. OConnor, D. Lyons, and G. M. Lyons. Accelerometers in rehabilitation medicine for older adults. *Age and Ageing*, 34:556–560, 2005.
- [4] Chong-Zhi Di, Ciprian M. Crainiceanu, Brian S. Caffo, and Naresh M. Punjabi. Multilevel functional principal component analysis. *Annals of Applied Statistics*, 3(1):458–488, 2008.
- [5] Clyde W. Yancy etal. 2013 accf/aha guideline for the management of heart failure:. *Circulation*, 128:e240–e327, 2013.

BIBLIOGRAPHY

- [6] Richard Everson and Stephen Roberts. Inferring the eigenvalues of covariance matrices from limited, noisy data. *IEEE Transactions on Signal Processing*, 48(7):2083–2091, 2000.
- [7] A. R. Feinstein, B. R. Josephy, and C. K. Wells. Scientific and clinical problems in indexes of functional disability. *Annals of Internal Medicine*, 105:413–420, 1986.
- [8] P. M. Grant, C. G. Ryan, W. W. Tigbe, and M. H. Granat. The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *British Journal of Sports Medicine*, 40:992–997, 2006.
- [9] Jean Hausser and Korbinian Strimmer. Entropy inference and the james-stein estimator with application to nonlinear gene association networks. *Journal of Machine Learning Research*, 10:1469–1484, 2009.
- [10] D. Hendelman, K. Miller, C. Baggett, E. Debold, and P. Freedson. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine & Science in Sports & Exercise*, 32:442–449, 2000.
- [11] Jason Howell, Michelle Strong, and etal. Maximum daily 6 minutes of activity: An index of functional capacity derived from actigraphy and its application to older adults with heart failure. *Journal of the American Geriatric Society*, 58:931–936, 2010.

BIBLIOGRAPHY

- [12] NL Johnson, HL Kirchner, and CL Rosen. Sleep estimation using wrist actigraphy in adolescents with and without sleep disordered breathing: a comparison of three data modes. *SLEEP*, 30:899–905, 2007.
- [13] S. Kozey-Keadle, A. Libertine, K. Lyden, J. Staudenmayer, and P.S. FREEDSON. Validation of wearable monitors for assessing sedentary behavior. *Medicine & Science in Sports & Exercise*, 43(8):1561, 2011.
- [14] I. McDowell and C. Newell. *Measuring Health: A Guide to Rating Scales and Questionnaires*. Oxford University Press, New York, 1987.
- [15] Thomas P. Minka. Automatic choice of dimensionality for pca. *Advances in Neural Information Processing Systems 13*, pages 598–604, 2000.
- [16] Philip R. Nader, Robert H. Bradley, Renate M. Houts, Susan L. McRitchie, and Marion OBrien. Moderate-to-vigorous physical activity from ages 9 to 15 years. *Journal of the American Medical Association*, 300:295–305, 2008.
- [17] Charles Pollak, Warren Tryon, Haikady Nagaraja, and Roger Dzwonczyk. How accurately does wrist actigraphy identify the states of sleep and wakefulness? *Sleep*, 24:957–965, 2001.
- [18] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman. Activity recognition from accelerometer data. In *Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence*, pages 1541–1546. AAAI Press, 2005.

BIBLIOGRAPHY

- [19] Troiano R.P., Berrigan D., Masse L.C., Tilert T., and McDowell M. Physical activity in the united states measured by accelerometer. *Medicine and Science in Sports and Exercise*, 40(1):181–188, 2008.
- [20] Pate R.R., Pratt M., and etal. Physical activity and public health. a recommendation from the centers for disease control and prevention and the american college of sports medicine. *Journal of the American Medical Association*, 273:402–407, 1995.
- [21] USDHSS. Chapter 22. physical activity and fitness. *Healthy People 2010*, pages 22.5–22.39, 2000.

Vita

Ximin Li, raised in Wuhan, China, graduated from the Ohio State University in with B.S in Math in 2012, and enrolled in the Biostatistics ScM program at Johns Hopkins School of Public Health. While at Johns Hopkins School of Public Health, her research focuses on physical activity over time recorded by accelerometer device. She is also a biostatistics consultant at Johns Hopkins Biostatistics center, consulting on biostatistical issues to fellow students and assist with study designs, statistical analysis and data interpretation in research project.